Report of ClimMob project Pot21A

18 February, 2021

Table of Contents

[Introduction 3](#_Toc64571113)

[ClimMob 3](#_Toc64571114)

[How to cite 4](#_Toc64571115)

[Section 1: Headline results 5](#_Toc64571116)

[Overall differences in rankings 8](#_Toc64571117)

[Correlation between the yield and the other traits 9](#_Toc64571118)

[Section 2: Yield, data summary and exploratory analysis 12](#_Toc64571119)

[Assessment of varieties 12](#_Toc64571120)

[Pairwise contests 13](#_Toc64571121)

[Plackett-Luce Model estimates 14](#_Toc64571122)

[Plackett-Luce Model with covariates for yield 18](#_Toc64571123)

[References 20](#_Toc64571124)

# Introduction

In agriculture, the local environmental conditions determine to a large degree which technological solutions are the most suitable. In dry soils, for example, drought-resistant crop varieties will outperform other varieties, but in wet soils these same varieties may do worse than most. Not only drought, but an entire range of problems including excessive heat, floods, new pests and diseases tend to intensify under climate change. This multitude of limiting factors requires multiple technological solutions, tested in diverse environments.

Citizen science is based on the cooperation of ‘citizen scientists’ or observers (paid or unpaid). Researchers assign small tasks (observations, experiments) that, once completed and gathered, contribute with a great amount of information to science. One of the advantages of citizen science is that agricultural researchers can get access to many environments by crowdsourcing their experiments. As farmers contribute with their time, skills and knowledge to the investigation, researchers are able to do more tests than in a traditional experimental design. Also citizen scientists acquire new knowledge, abilities and information useful for future challenges of their work.

## ClimMob

The primary goal of ClimMob is to support the selection of innovative technologies (*e.g.* new crop varieties, management, new product). ClimMob serves to prepare and analyze citizen science experiments in which a large number of participants observe and compare different technological options under a wide range of conditions (*1*).

ClimMob software assigns a limited number of items (typically 3) to each participant, who will compare their performance. Each participant gets a different combination of items drawn from a much larger set of items (typically 15-25). Comparisons of this kind are thought to be a very reliable way to obtain data. Once the results of the small tasks have been collected, ClimMob builds an image of the whole set of assigned objects, combining all observations. ClimMob not only reconstructs the overall ordering of items, but also takes into account differences and similarities between participants and the conditions under which they observe (e.g. socio-economic and plot environmental traits). It assigns similar participants to groups that each corresponds among different group profiles. Groups are created on the basis of whichever items which have been collected, that are found to be significantly linked to the observed rankings.

ClimMob uses Plackett-Luce models to analyze ranking data with the R (*2*) package ‘PlackettLuce’ (*3*). It automatically generates analytical reports, as well as individualized information sheets for each participant using the R packages ‘knitr’ (*4*) and ‘rmarkdown’ (*5*). Organizing the data relies on packages ‘ClimMobTools’ (*6*), ‘gosset’ (*7*), ‘gtools’ (*8*), ‘jsonlite’ (*9*), ‘partykit’ (*10*), ‘psychotools’ (*11*) and ‘qvcalc’ (*12*). Summaries and data visualization are supported by packages ‘igraph’ (*13*), ‘ggparty’ (*14*), ‘ggplot2’ (*15*), ‘ggrepel’ (*16*), ‘gridExtra’ (*17*) ‘leaflet’ (*18*), ‘multcompView’ (*19*), ‘patchwork’ (*20*), ‘png’ (*21*), ‘plotrix’ (*22*) and ‘pls’ (*23*). The workflow used to analyze the data and produce this report is documented in Zenodo (*24*).

## How to cite

If you publish any results generated with ClimMob, you should cite a number of articles as the package builds on various contributions. Van Etten et al. (2019) (*1*) introduced the crowdsourcing philosophy behind ClimMob. It is important to mention that ClimMob is implemented in R, a free, open-source analysis software (*2*). Methodologically, if you report on the Plackett-Luce tree results, you should mentioned that ClimMob applies the Plackett-Luce model published by Turner et al. 2020 (*3*). To cite ClimMob itself, mention van Etten et al. (2020) (*25*).

# Section 1: Headline results

Overall there were 150 farmers registered to this project. Each farmer assessed 3 different varieties and ranked them in order of its ‘yield’. In addition they also provided rankings for 10 additional trait(s):

Table 1.1. Summary of traits assessed in this project and valid answers used in this report.

|  |  |  |  |
| --- | --- | --- | --- |
| Trait | Data collection moment | Question asked | Number of valid answers |
| Plant vigour | Vegetative 1 (Day 45) | Which is the best growing variety?, Which is the worst growing variety? | 148 |
| Bacterial wilt damage | Vegetative 1 (Day 45) | Which variety shows the least damage from bacterial wilt?, Which variety shows the most damage from bacterial wilt? | 38 |
| Disease and insect damage | Vegetative 1 (Day 45) | Which variety shows the least damage from other diseases or pests?, Which variety shows the most damage from other diseases or pests? | 24 |
| Bacterial wilt damage | Vegetative 2 (Day 75) | Which variety shows the least damage from bacterial wilt?, Which variety shows the most damage from bacterial wilt? | 150 |
| Disease and insect damage | Vegetative 2 (Day 75) | Which variety shows the least damage from other diseases or pests?, Which variety shows the most damage from other diseases or pests? | 150 |
| Maturity | Post-harvest 1 (5 days after harvest) | Which variety is ready for harvesting the quickest?, Which variety is ready for harvesting the slowest? | 126 |
| Yield | Post-harvest 1 (5 days after harvest) | Which variety gives the best yield?, Which variety gives the worst yield? | 126 |
| Tuber size | Post-harvest 1 (5 days after harvest) | Which variety produces the largest tubers?, Which variety produces the smallest tubers? | 126 |
| Tuber appearance | Post-harvest 1 (5 days after harvest) | Which variety has the best appearing tubers?, Which variety has the worst appearing tubers? | 126 |
| Marketability at harvest | Post-harvest 1 (5 days after harvest) | Which variety is the easiest to sell at harvest?, Which variety is the most difficult to sell at harvest? | 126 |
| Taste | Post-harvest 1 (5 days after harvest) | Which variety tastes best after cooking?, Which variety tastes worst after cooking? | 126 |

The map below shows the distribution of the trials in this project. If you registered more than one GPS location per farmer the map shows the GPS registry which had the largest number of valid GPS records. We used the GPS points from ‘registration\_REG\_farm\_geogitude’. To respect farmer’s privacy, the coordinates were clustered with a 0.05 arc-degree resolution. You can find the original coordinates in your ClimMob data.

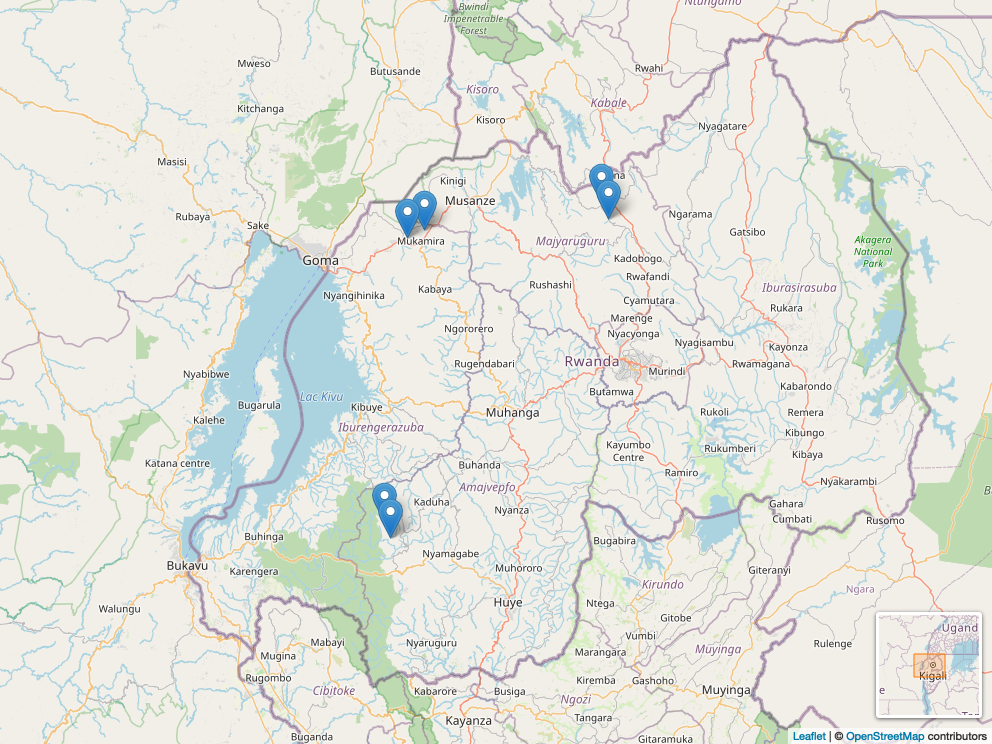


Table 1.2 shows the 11 varieties assessed within this project, with the frequency and percentage of farmers who assessed each variety. If the varieties have large names (> 10 characters) an abbreviation was applied across the figures in this report.

Table 1.2. Frequency of varieties assessed.

|  |  |  |  |
| --- | --- | --- | --- |
| Variety | Abbreviation | Freq | Relative freq |
| Cruza | Cruza | 42 | 28% |
| Gisubizo | Gisubizo | 41 | 27.3% |
| Izihirwe | Izihirwe | 40 | 26.7% |
| Jyambere | Jyambere | 41 | 27.3% |
| Kazeneza | Kazeneza | 41 | 27.3% |
| Kirundo | Kirundo | 41 | 27.3% |
| Ndamira | Ndamira | 41 | 27.3% |
| Ndeze | Ndeze | 41 | 27.3% |
| Nkunganire | Nkng | 40 | 26.7% |
| Seka | Seka | 41 | 27.3% |
| Twihaze | Twihaze | 41 | 27.3% |

Figure 1.1 shows that the 11 varieties are all connected to each other. That means that they all co-occurred at least once in an incomplete block of 3 varieties.

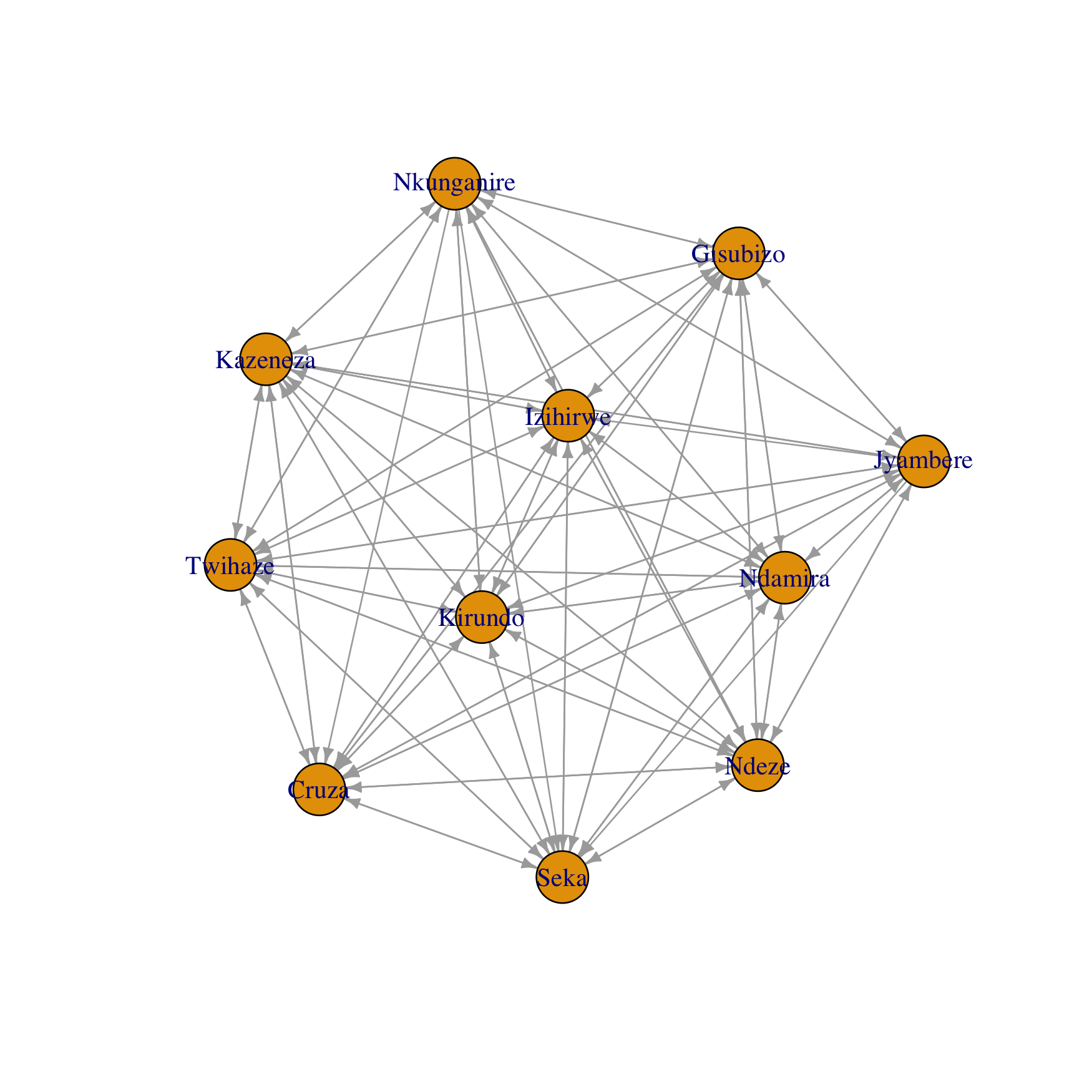


Figure 1.1. Network representation of varieties tested in this project.

## Overall differences in rankings

This reports takes the trait ‘yield’ as the reference trait for the analysis. There were statistically significant differences found in the rankings of varieties in the trait ‘yield’ (p = 1.24e-04). The best ranked varieties overall were Cruza, Seka, Kirundo. Statistically significant differences were found in the trait(s) ‘plant vigour [vegetative 1 (day 45)]’, ‘bacterial wilt damage [vegetative 2 (day 75)]’, ‘maturity [post-harvest 1 (5 days after harvest)]’, ‘tuber size [post-harvest 1 (5 days after harvest)]’.

A summary of the p-values testing the hypothesis that there were differences in the rankings for each trait, and the list of varieties which were significantly highest and lowest ranked overall, are summarised in Table 1.3.

Table 1.3. Summary of differences found in varieties by trait.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Trait | Data collection moment | Best Ranked | Worst Ranked | p.value |  |
| Plant vigour | Vegetative 1 (Day 45) | Cruza, Seka, Kirundo | Ndeze, Jyambere, Nkunganire | 1.24e-04 | \*\*\* |
| Bacterial wilt damage | Vegetative 1 (Day 45) | Kazeneza, Twihaze | Jyambere, Ndeze, Nkunganire | 8.53e-13 | \*\*\* |
| Disease and insect damage | Vegetative 1 (Day 45) | No significant difference | No significant difference | 3.36e-01 |  |
| Bacterial wilt damage | Vegetative 2 (Day 75) | No significant difference | No significant difference | 4.36e-01 |  |
| Disease and insect damage | Vegetative 2 (Day 75) | Izihirwe, Jyambere, Cruza | Gisubizo, Kirundo, Twihaze | 9.16e-03 | \*\* |
| Maturity | Post-harvest 1 (5 days after harvest) | No significant difference | No significant difference | 4.73e-01 |  |
| Yield | Post-harvest 1 (5 days after harvest) | Ndamira, Kirundo, Seka | Nkunganire, Jyambere, Cruza | 1.54e-09 | \*\*\* |
| Tuber size | Post-harvest 1 (5 days after harvest) | Seka, Twihaze, Kirundo | Ndeze, Nkunganire, Cruza | 5.19e-09 | \*\*\* |
| Tuber appearance | Post-harvest 1 (5 days after harvest) | No significant difference | No significant difference | 7.43e-01 |  |
| Marketability at harvest | Post-harvest 1 (5 days after harvest) | No significant difference | No significant difference | 4.43e-01 |  |
| Taste | Post-harvest 1 (5 days after harvest) | No significant difference | No significant difference | 4.27e-01 |  |

## Correlation between the yield and the other traits

Table 1.4 shows, for each trait evaluated in the project, the frequency for which the rankings matched with the trait ‘yield’. Best and worst agreement represents the percentage for which the best and worst variety for the trait matched the best and worst ‘yield’. Complete ranking agreement shows the proportion of correlation on the full ranking with the trait ‘Yield’ as baseline using the Kendall correlation coefficient (*26*).

Table 1.4. Correlation between ‘yield’ and the others traits assessed in this project.

|  |  |  |  |
| --- | --- | --- | --- |
| Trait | Complete Ranking Agreement (Kendall tau) | Agreement with best | Agreement with worst |
| Plant vigour [Vegetative 1 (Day 45)] | 25% | 50% | 37.5% |
| Bacterial wilt damage [Vegetative 1 (Day 45)] | 50% | 62.5% | 75% |
| Disease and insect damage [Vegetative 1 (Day 45)] | 8.3% | 37.5% | 25% |
| Bacterial wilt damage [Vegetative 2 (Day 75)] | -8.3% | 0% | 62.5% |
| Disease and insect damage [Vegetative 2 (Day 75)] | 0% | 25% | 37.5% |
| Maturity [Post-harvest 1 (5 days after harvest)] | 0% | 37.5% | 25% |
| Tuber size [Post-harvest 1 (5 days after harvest)] | 8.3% | 37.5% | 50% |
| Tuber appearance [Post-harvest 1 (5 days after harvest)] | 58.3% | 62.5% | 75% |
| Marketability at harvest [Post-harvest 1 (5 days after harvest)] | 58.3% | 75% | 62.5% |
| Taste [Post-harvest 1 (5 days after harvest)] | 41.7% | 62.5% | 50% |

The trait which had the strongest relationship with ‘yield’ was ‘tuber appearance [post-harvest 1 (5 days after harvest)]’. Overall, the rankings for ‘tuber appearance [post-harvest 1 (5 days after harvest)]’ matched the rankings for ‘yield’ 58% of the time. The trait which had the weakest relationship with ‘yield’ was ‘bacterial wilt damage [vegetative 2 (day 75)]’. Overall the rankings for ‘bacterial wilt damage [vegetative 2 (day 75)]’ matched the rankings for ‘yield’ -8% of the time. The correlation with the other traits is also shown in Figure 1.2.

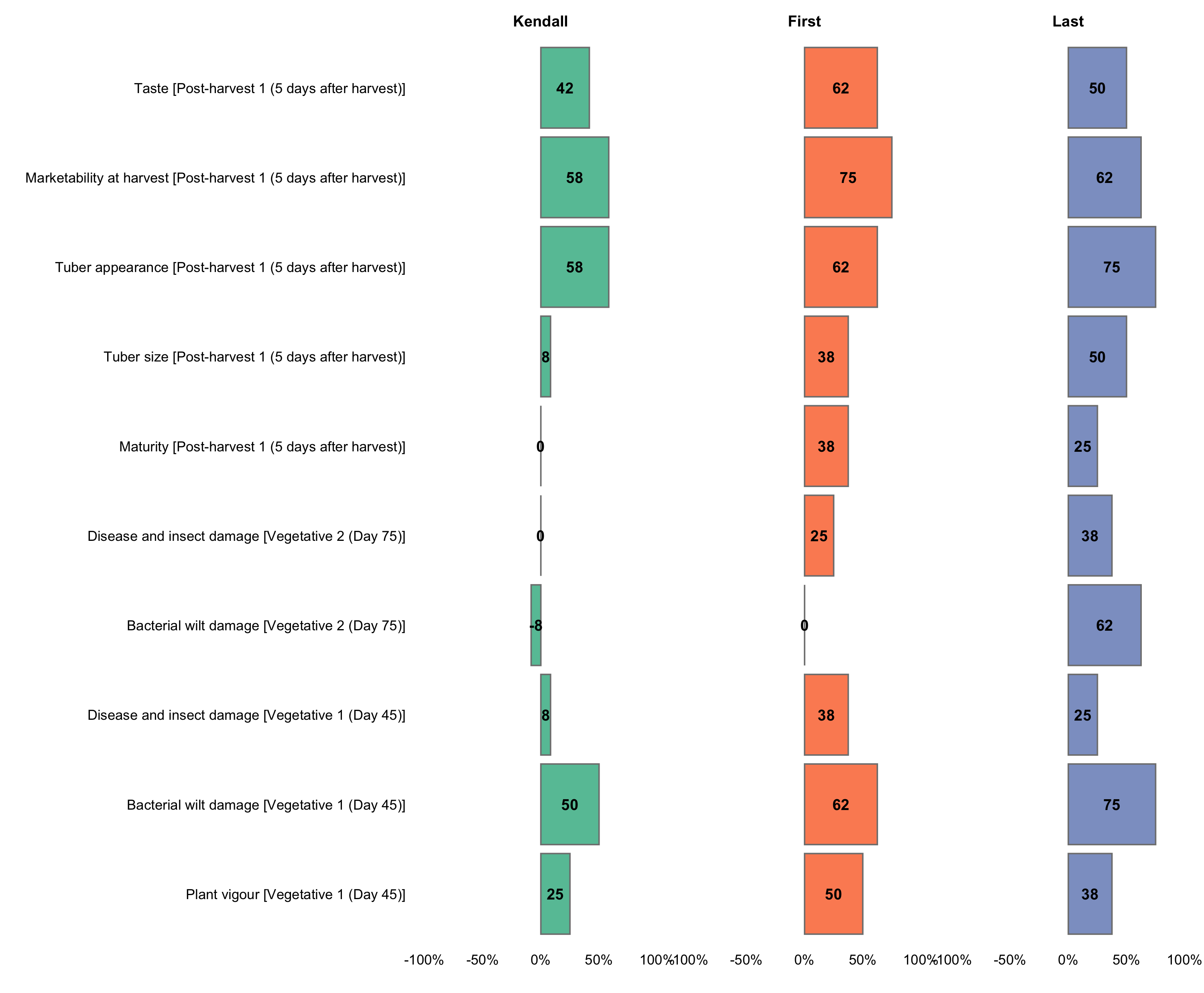


Figure 1.2. Correlation between individual traits and yield.

Partial least squares regression was used to determine relationship between the other traits and ‘yield’. The first two components recombining the specific traits are able to explain the variability in ‘yield’. The dashed line represents the ‘yield’ with an increase in performance as the x and y increase. The variety positioned close to the dashed line will be performing equally across all traits. The variety positioned further away from the dashed line, on either side, will have varying performance in different traits. Better performance in a given trait will correspond with arrows pointing in the direction away from the dashed line and worse performance in traits directed on the opposite side.

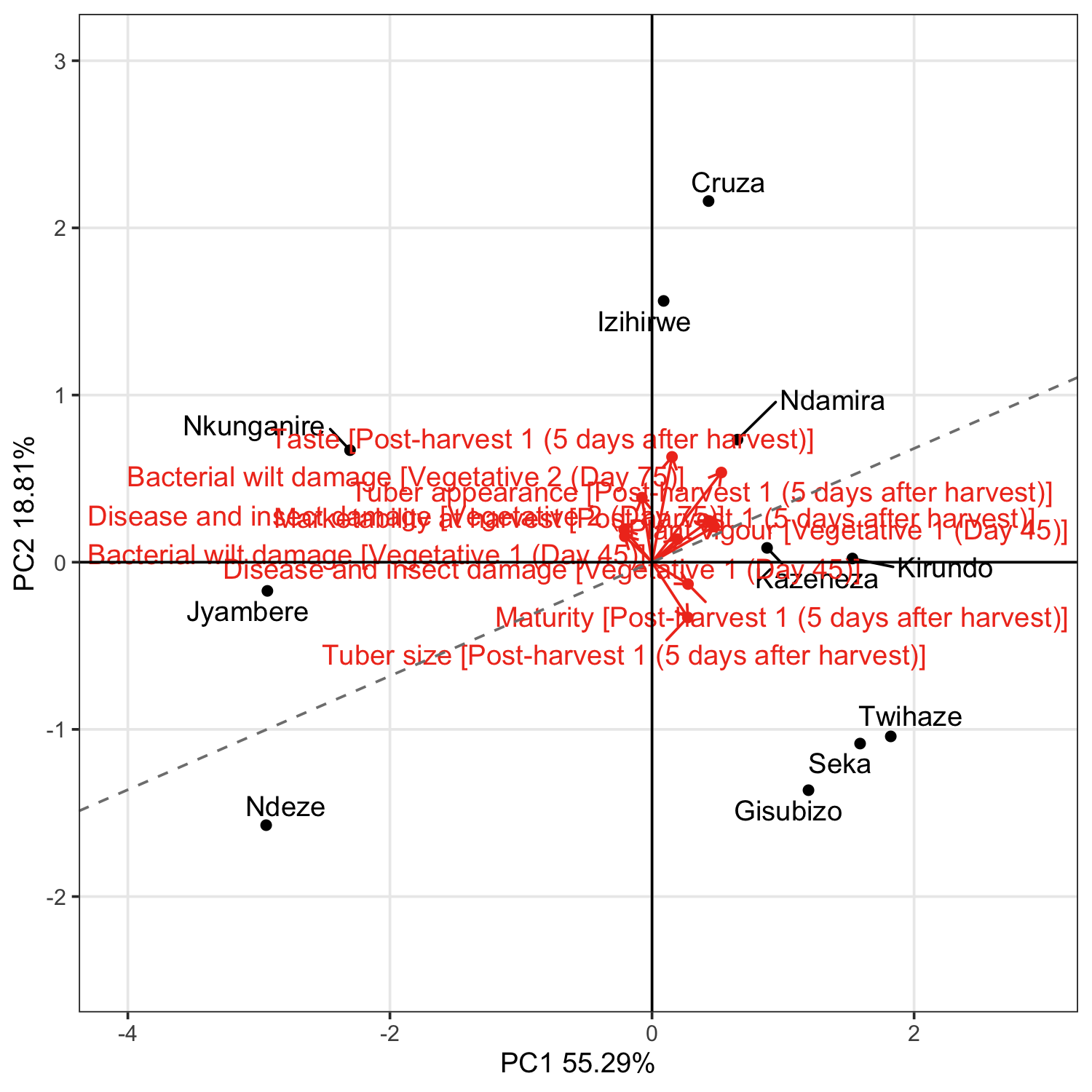


Figure 1.3. Partial least squares biplot of relationship between the tested varieties and the traits evaluated.

# Section 2: Yield, data summary and exploratory analysis

## Assessment of varieties

Exploratory analysis within the following section summarises results from the rankings on ‘yield’. Given the structure of a ClimMob trail, where each farmer only assessed 3 of the 11 possible varieties these results may be skewed if certain varieties were randomly assigned to face worse varieties than others. This is particularly a potential issue within a smaller trial, as due to the randomization process the potential for an unbalanced assignment decreases as the sample size increases. Results from other sections, and in the overall summary use Plackett-Luce models (*3*), to adjust for any imbalance. Performance of each of the varieties for ‘yield’ is summarised in Table 2.1.

Table 2.1. Favourability scores for ‘yield’.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variety | N | Top Ranked | Bottom Ranked | Net Favourability Score |
| Cruza | 34 | 58.8% | 11.8% | 47.1 |
| Seka | 34 | 44.1% | 17.6% | 26.5 |
| Kirundo | 36 | 50% | 27.8% | 22.2 |
| Ndamira | 37 | 43.2% | 27% | 16.2 |
| Kazeneza | 32 | 40.6% | 28.1% | 12.5 |
| Twihaze | 36 | 30.6% | 38.9% | -8.3 |
| Gisubizo | 33 | 24.2% | 33.3% | -9.1 |
| Izihirwe | 38 | 21.1% | 31.6% | -10.5 |
| Nkunganire | 29 | 13.8% | 44.8% | -31.0 |
| Jyambere | 33 | 15.2% | 48.5% | -33.3 |
| Ndeze | 36 | 22.2% | 58.3% | -36.1 |

This shows the percentage of farmers who assessed the varieties as the best among the 3 varieties they were provided, the percentage of farmers who included the variety as their worst, the percentage of ‘head to head contests’ for which the variety won and the net *favourability* score (Figure 2.1). A score of +100 indicates the variety won all ‘contests’ it was involved in, a score of 0 indicates an equal number of wins and losses, a score of -100 indicates the variety lost all contests.

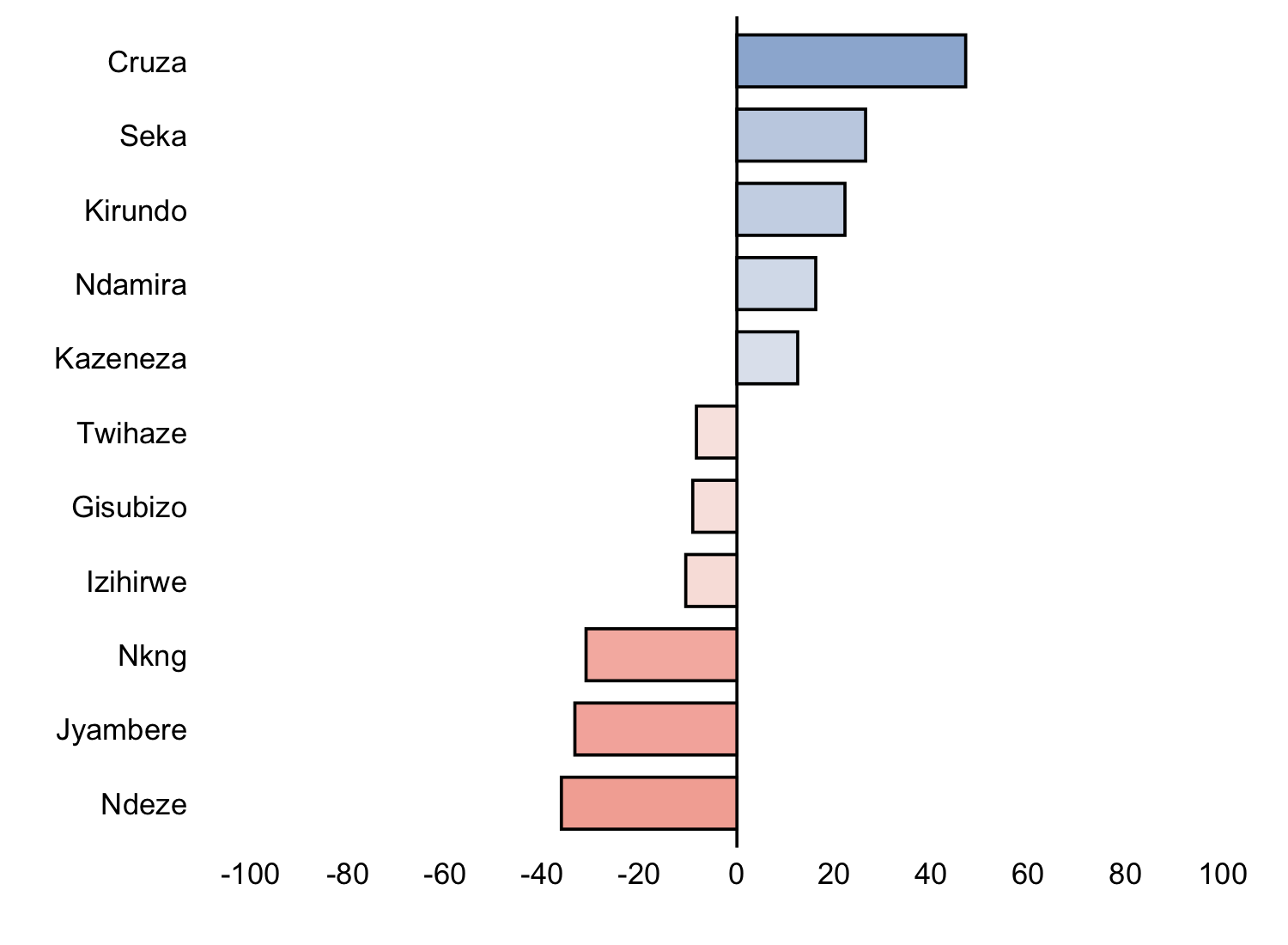


Figure 2.1. Net favourability scores for ‘yield’.

The variety Cruza was the ‘best’ variety for the trait yield being ranked highest by 58.8% of the 150 farmers who assessed this variety .

## Pairwise contests

Figure 2.3 shows the outcomes of all pairwise contests between the varieties included in the project for ‘yield’. Each panel shows the performance of one variety against all the other varieties, and shows the percentage of the times in which the panelled variety was ranked above the other varieties shown as bars.

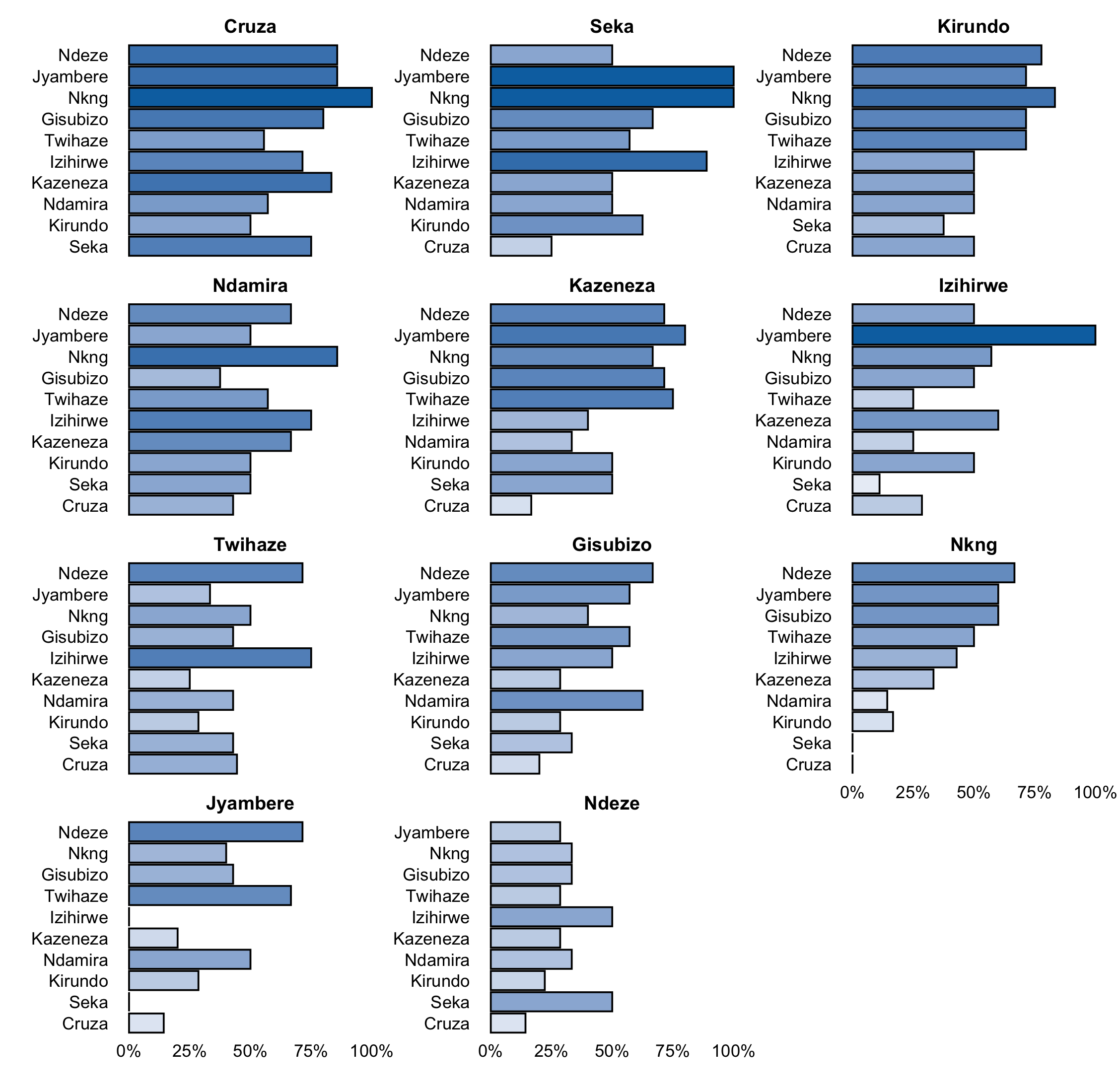


Figure 2.3. Head to head performance for ‘yield’.

## Plackett-Luce Model estimates

Table 2.2 shows the results from the likelihood ratio test from the Plackett-Luce model for the trait ‘yield’ of the different varieties. The hypothesis being tested is that there is no difference in the assessments of any of the different varieties.

Table 2.2. Likelihood ratio test results from fitted Plackett-Luce model with rankings from ‘yield’.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| model | logLikelihood | DF | Statistic | Pr(>Chisq) |
| NULL | -225.7617 | 378 | NA | NA |
| Yield | -208.2564 | 368 | 35.01054 | 1.24e-04 \*\*\* |

Figure 2.4 shows the estimates of the model coefficients with 84% confidence intervals. The purpose of this graph is to be able to best distinguish between the relative strength of each of the varieties assessed. As such the coefficient estimates themselves are not directly interpretable, but it can be concluded that a higher value for the coefficient indicates that a variety has been ranked as best more often. The 84% confidence width is chosen so that non-overlapping confidence intervals could be interpreted as indicating significant differences at the alpha = 0.1. This may not match exactly with the mean separation groupings, as these groupings also take into account multiple testing through the Benjamini and Hochberg adjustment (*27*).

Mean separation analysis was also conducted to indicate, using letters, which varieties are significantly more preferred than others in terms of ‘yield’. When varieties have at least one letter in common, there is not enough evidence from the experiment to be confident about their relative order at the alpha = 0.1.

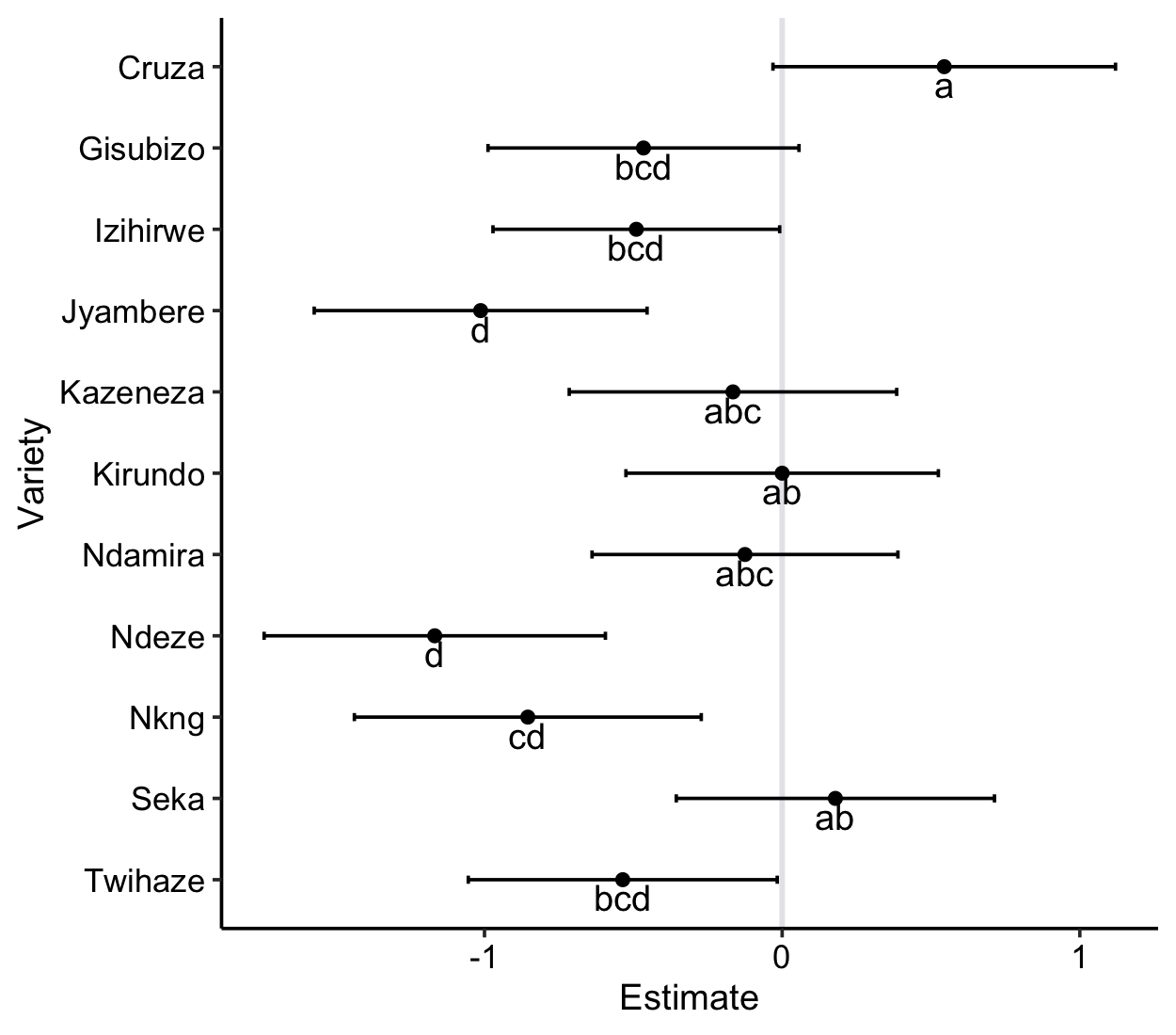


Figure 2.4. Plackett-Luce Model estimates (log-worth) of tested varieties for ‘yield’ with 84% confidence intervals. Different letters indicate significant differences at p < 0.1. The variety Kirundo is set as reference (log-worth arbitrarily set to zero).

The same information in Figure 2.4 is shown in Table 2.3 below.

Table 2.3. Plackett-Luce Model estimates (log-worth) of tested varieties for ‘yield’ with 84% confidence intervals. Different letters indicate significant differences at p < 0.1. The variety Kirundo is set as reference (log-worth arbitrarily set to zero).

|  |  |  |  |
| --- | --- | --- | --- |
| Variety | Estimate | quasiSE | Group |
| 1 | 0.5445 | 0.2936 | a |
| 2 | 0.1788 | 0.2727 | ab |
| 3 | 0.0000 | 0.2678 | ab |
| 4 | -0.1249 | 0.2621 | abc |
| 5 | -0.1652 | 0.2807 | abc |
| 6 | -0.4658 | 0.2665 | bcd |
| 7 | -0.4899 | 0.2457 | bcd |
| 8 | -0.5355 | 0.2648 | bcd |
| 9 | -0.8545 | 0.2973 | cd |
| 10 | -1.0130 | 0.2853 | d |
| 11 | -1.1671 | 0.2925 | d |

Table 2.4 and Figure 2.5 use the coefficients from the Plackett-Luce model to estimate the probability of each variety being considered to be the top ranked variety in a direct comparison between all of the possible varieties.

Table 2.4. Percentage probability of being the best ranked for ‘yield’.

|  |  |
| --- | --- |
| Variety | Win probability |
| Cruza | 20.1% |
| Seka | 13.9% |
| Kirundo | 11.7% |
| Ndamira | 10.3% |
| Kazeneza | 9.9% |
| Gisubizo | 7.3% |
| Izihirwe | 7.1% |
| Twihaze | 6.8% |
| Nkunganire | 5% |
| Jyambere | 4.2% |
| Ndeze | 3.6% |

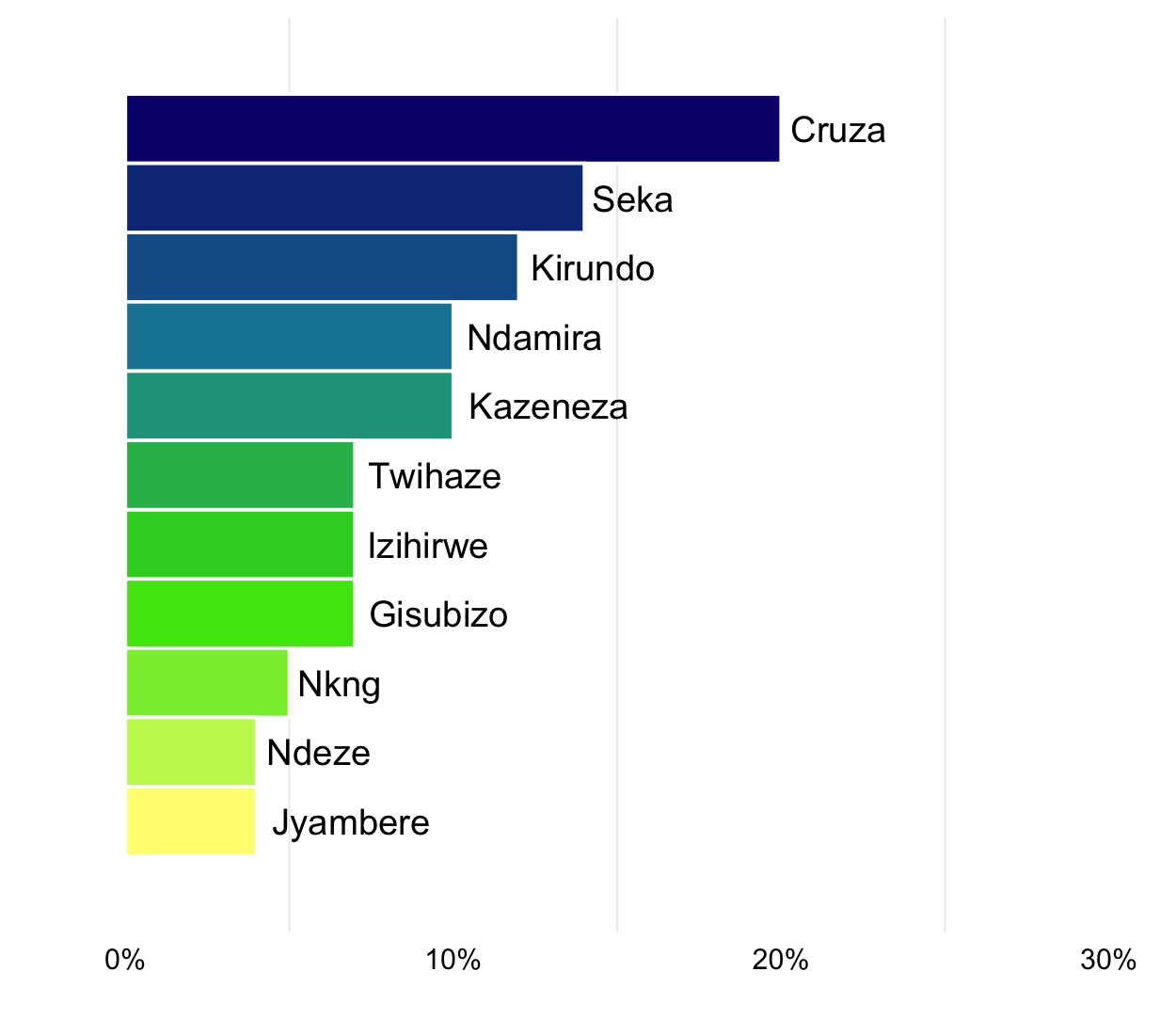


Figure 2.5. Probability of being the best ranked for ‘yield’.

## Plackett-Luce Model with covariates for yield

A model-based recursive partitioning method (*28*) was used to determine which of the covatiates, if any, had significant relationships with the rankings of yield. This approach identifies sub-groups in the data for which the rankings of the different varieties are significantly different to each other. This analysis required that all rankings and covariates are completely available for the entire dataset. Here we used 83 out of 126 valid rankings for yield.

None of the covariates tested were found to have a statistically significant relationship to ‘yield’ with an alpha = 0.1.

Table 2.7. Univariate p-values for first split in Plackett-Luce tree model for the trait ‘yield’.

|  |  |  |
| --- | --- | --- |
| Covariate | Question | p.value |
| Consent | Do you consent to participating in this trial and sharing us your trial information? | NA |
| TrialSuccess | Did the trial continue until this time? | 9.94e-01 |
| TrialSuccess1 | Did the trial continue until this time? | 8.92e-01 |
| TrialSuccess2 | Did the trial continue until this time? | 8.38e-01 |
| TubersToObserve | Does farmer still have enough tubers of the 3 varieties to enable comparison? | 2.52e-01 |

# References

1. J. van Etten, E. Beza, L. Calderer, K. Van Duijvendijk, C. Fadda, B. Fantahun, Y. G. Kidane, J. van de Gevel, A. Gupta, D. K. Mengistu, D. Kiambi, P. N. Mathur, L. Mercado, S. Mittra, M. J. Mollel, J. C. Rosas, J. Steinke, J. G. Suchini, K. S. Zimmerer, First experiences with a novel farmer citizen science approach: crowdsourcing participatory variety selection through on-farm triadic comparisons of technologies (tricot). *Experimental Agriculture*. **55**, 275–296 (2019).

2. R Core Team, R: A language and environment for statistical computing. version 3.6.2. (2019), (available at <https://r-project.org/>).

3. H. L. Turner, J. van Etten, D. Firth, I. Kosmidis, Modelling rankings in R: the PlackettLuce package. *Computational Statistics* (2020), doi:[10.1007/s00180-020-00959-3](https://doi.org/10.1007/s00180-020-00959-3).

4. Y. Xie, *Dynamic Documents with R and knitr* (Chapman; Hall/CRC, Boca Raton, Florida, ed. 2nd, 2015; <https://yihui.org/knitr/>).

5. Y. Xie, J. J. Allaire, G. Grolemund, *R Markdown: The Definitive Guide* (Chapman; Hall/CRC, Boca Raton, Florida, 2018; <https://bookdown.org/yihui/rmarkdown>).

6. K. de Sousa, J. van Etten, ClimMobTools: API Client for the ’ClimMob’ platform (2020), (available at <https://CRAN.R-project.org/package=ClimMobTools>).

7. K. de Sousa, J. van Etten, S. Dumble, N. Greliche, J. Steinke, gosset: Modelling Metadata and Crowdsourced Citizen Science (2020), (available at <https://agrobioinfoservices.github.io/gosset/>).

8. G. R. Warnes, B. Bolker, T. Lumley, gtools: Various R Programming Tools (2018), (available at <https://CRAN.R-project.org/package=gtools>).

9. J. Ooms, The jsonlite Package: A Practical and Consistent Mapping Between JSON Data and R Objects. *arXiv:1403.2805 [stat.CO]* (2014) (available at <https://arxiv.org/abs/1403.2805>).

10. A. Zeileis, T. Hothorn, K. Hornik, Model-Based Recursive Partitioning. *Journal of Computational and Graphical Statistics*. **17**, 492–514 (2008).

11. J. Kopf, C. Strobl, A. Zeileis, Anchor Selection Strategies for DIF Analysis: Review, Assessment, and New Approaches. *Educational and Psychological Measurement*. **75**, 22–56 (2015).

12. D. Firth, qvcalc: Quasi Variances for Factor Effects in Statistical Models (2020), (available at <https://CRAN.R-project.org/package=qvcalc>).

13. G. Csardi, T. Nepusz, The igraph software package for complex network research. *InterJournal*. **Complex Systems**, 1695 (2006).

14. M. Borkovec, N. Madin, *ggparty: ’ggplot’ Visualizations for the ’partykit’ Package* (2019; <https://CRAN.R-project.org/package=ggparty>).

15. H. Wickham, *ggplot2: Elegant Graphics for Data Analysis* (Springer-Verlag New York, 2016; <https://ggplot2.tidyverse.org>).

16. K. Slowikowski, *ggrepel: Automatically Position Non-Overlapping Text Labels with ’ggplot2’* (2020; <https://CRAN.R-project.org/package=ggrepel>).

17. B. Auguie, *gridExtra: Miscellaneous Functions for "Grid" Graphics* (2017; <https://CRAN.R-project.org/package=gridExtra>).

18. J. Cheng, B. Karambelkar, Y. Xie, *leaflet: Create Interactive Web Maps with the JavaScript ’Leaflet’ Library* (2019; <https://CRAN.R-project.org/package=leaflet>).

19. S. Graves, H.-P. Piepho, L. S. with help from Sundar Dorai-Raj, multcompView: Visualizations of Paired Comparisons (2019), (available at <https://CRAN.R-project.org/package=multcompView>).

20. T. L. Pedersen, *patchwork: The Composer of Plots* (2019; <https://CRAN.R-project.org/package=patchwork>).

21. S. Urbanek, *png: Read and write PNG images* (2013; <https://CRAN.R-project.org/package=png>).

22. L. J, Plotrix: a package in the red light district of R. *R-News*. **6**, 8–12 (2006).

23. B.-H. Mevik, R. Wehrens, K. H. Liland, pls: Partial Least Squares and Principal Component Regression (2019), (available at <https://CRAN.R-project.org/package=pls>).

24. K. de Sousa, S. Dumble, N. Greliche, B. Madriz, J. van Etten, *Workflow for data analysis and report with crowdsourced citizen science-generated data* (Zenodo, 2020; <https://doi.org/10.5281/zenodo.3976631>).

25. J. van Etten, R. Manners, J. Steinke, E. Matthus, K. de Sousa, *The tricot approach: Guide for large-scale participatory experiments* (Alliance of Bioversity International; CIAT; Alliance Bioversity International; CIAT, Rome, Italy, 2020; <https://hdl.handle.net/10568/109942>).

26. M. G. Kendall, A new measure of ranking correlation. *Biometrika*. **30**, 81–93 (1938).

27. Y. Benjamini, Y. Hochberg, Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)*. **57**, 289–300 (1995).

28. C. Strobl, J. Malley, G. Tutz, An introduction to recursive partitioning: rationale, application, and characteristics of classification and regression trees, bagging, and random forests. *Psychological methods*. **14**, 323–48 (2009).